DOCUMENT RESUME

ED 408 348 TM 026 638

AUTHOR Svensson, Allan; And Others
TITLE Changes in Intelligence from 1960 to 1995 in Relation to Cohort, Gender, and Socioeconomic Background.
PUB DATE Mar 97
NOTE 29p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, March 24-28, 1997).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Change; Cohort Analysis; Educational History; Elementary Education; *Elementary School Students; Foreign Countries; Intelligence; *Intelligence Tests; Preschool Education; Sampling; Scores; Sex Differences; Socioeconomic Status; Spatial Ability; *Test Results; Thinking Skills; *Verbal Ability; Vocabulary
IDENTIFIERS Reasoning Tests; *Sweden

ABSTRACT

As part of a continuing evaluation of the Swedish school system, changes in intelligence test scores from 1960 to 1995 were studied in relation to cohort, gender, and socioeconomic background. Changes in verbal, spatial, and reasoning ability were studied by comparing results on identical tests given to representative samples of Swedish 13-year-olds on six occasions from 1961 to 1995. The first data collection involved about 12,000 students, 90% of whom were in sixth grade in 1961. In 1966 approximately 10,500 students born in 1953 were studied. The 1980 data collection, the third overall, used approximately 9,000 students, mostly born in 1967. The fourth, fifth, and sixth data samples were drawn in about the same way as the third sample, using mostly students born in 1972, 1977, and 1982. Verbal test scores tended to rise until 1980, after which there was a stronger tendency in the opposite direction. However, behind the decrease are changes in the proportion of older Swedish and foreign words, so that it is really only possible to say that students' vocabularies are different than they were 35 years earlier. Results in the spatial test rose considerably up to 1995 and girls' results more nearly approached boys' over the period. These changes may reflect the introduction of preschool and comprehensive school education in Sweden. In the reasoning test, increases were found for the first two decades, again perhaps reflecting changes in the educational system. An appendix contains five tables of test results. (Contains 3 figures, 16 tables, and 16 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Paper presented at the AERA Annual Meeting, Chicago, 24-28 March, 1997.

ED 408 348

CHANGES IN INTELLIGENCE FROM 1960 TO 1995 IN RELATION TO COHORT, GENDER, AND SOCIOECONOMIC BACKGROUND

Allan Svensson, Ingemar Emanuelsson & Sven-Eric Reuterberg.
Department of Education, Göteborg University.

INTRODUCTION

Tests for estimating mental ability have been in use since the beginning of the 20th century, and since then, there has been a lively discussion about intelligence development. Among other things, the question of whether the level of intelligence among children and youth is falling or rising, has been continuously discussed since the 1930's. As it is impossible to go into details here from this debate, we refer to reviews presented by Flynn (1987), Lynn (1990), Husén & Tuijnman (1991) and Brody (1992). However, we would like to point out one important reason for this long debate. It is extremely difficult to get empirical evidence about the real course of the development of intelligence. To acquire such evidence, you must have access to test results from large and representative samples who have been tested at the same age and with identical tests on different occasions. These conditions are very rarely met (see Halpern, 1992, p 91).

In spite of the fact that the requirements mentioned have almost never been fulfilled, many investigations are reported on attempts to clarify the general trends in changing levels of intelligence over time. A closer look at these studies will show that the results differ, but most of them indicate a slow but yet increasing average of mental ability up to the 1960's. Better and longer education, an extended availability of cultural activities, and a rising level of living standards are proposed as explanatory factors.

However, a decreasing level of mental ability since the 70's has been reported from the USA, especially in verbal-educational tests (Flanagan, 1976; Harnischfeger & Wiley, 1976). A similar trend in verbal ability has been observed in Sweden too (Emanuelsson, Reuterberg & Svensson, 1993), but the most striking outcome from this investigation - the results of which are incorporated in the subsequent analyses in this paper - is the extent to which the mean level of spatial and reasoning intelligence has risen between 1960 and 1980.

TM026638

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Allan Svensson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

BEST COPY AVAILABLE

A weak decreasing trend in verbal tests, and a strong increasing trend in non-verbal tests, are also found in a study published by Flynn (1987). This study is based on changes in results from 1950 to 1980 on various kinds of intelligence tests in fourteen countries, in Europe, North America, Asia, and New Zealand.

After a careful examination of his results, Flynn reaches the conclusion that the level of IQ, and particularly problem-solving ability, has increased massively since 1950. However, he does not give any detailed information on the causal factors behind these gains in intelligence. All he says is: *"Environmental factors with a large impact on IQ have not been identified"* (Flynn, 1987, p 190).

The reason why scores on verbal tests do not show the same rising trend as the scores on other tests is explained by the fact that verbal scores are more dependent on educational conditions, and that the quality of school education has declined.(op. cit., 184-185).

Flynn's conclusions are questioned by Lynn (1990), who argues that most of the increases in intelligence are caused by improvements in nutrition. As a contributory explanation of the relative weak development in verbal-educational tests, Lynn calls attention to the fact that school curricula have changed over time, and the verbal abilities measured by the tests were taught more thoroughly in earlier periods.

Husén & Tuijnman (1991) also critizise Flynn for not elaborating the importance of the expansion of formal schooling, in most industrial societies, as a possible explanation. By using the LISREL method, they demonstrate that schooling has a direct effect on adult IQ, even after variations in home background and child IQ are being controlled for. This finding is in agreement with results earlier presented by Härnqvist (1968). His results show that whilst child IQ has an effect on schooling outcomes, schooling per se has a substantial effect on IQ test scores. The same opinion is stressed by Brody (1992):

> *"Intelligence and educational accomplishment are linked and may be subject to bi-directional influences. Intelligence test scores predict performance in academic settings. The knowledge that is acquired in academic settings influences one's ability to solve problems and one's general intelligence. If the knowledge that is acquired in school influences scores on tests of intelligence, it is reasonable to assume that variations in the quantity and quality of education relate to performance on tests of intelligence."* (p 186).

According to our opinion, an individual's score on an intelligence test and changes over time depend on four kinds of factors:

1. *Biological factors*, for instance, hereditary factors, physical development and nutrition.

2. *Social factors*, of which home background and level of education are some of the most important.

3. *Error factors*. Under this category belong not only random errors of measurement, (unreliability) but also systematic errors such as effects of coaching and test-wiseness.

4. *Test content*. Over a period of 35 years it is reasonable to expect changes in test content, particularly for the verbal test due to changes in the every day language.

All the factors differ in significance from one test to another, and only one of the categories, the error factors, is possible to control and - at least to some extent - to eliminate, by using carefully constructed tests, tests with high reliability and by guarding the tests against unwarranted use.

The aim of the present investigation, is to study changes in verbal, spatial and reasoning ability, from the beginning of the 60's to 1995. We will especially study what changes have occurred concerning differences between boys and girls, and between students from various social groups. This will be done by comparing results on identical tests given to representative samples of Swedish 13-year old students on six occasions from 1961 to 1995.

SAMPLES AND VARIABLES

Samples

This study is a part of a longitudinal research project called "Evaluation Through Follow-up". The general aim of the project is a continuous evaluation of the Swedish school system. More specifically the main purposes are:

A. To make follow-up studies of large and nationally representative samples of students possible, to ascertain in what way geographic, social and psychological factors affect educational and vocational careers and to discover what changes the Swedish educational reforms have brought about in these respects.

B. To provide a basis for studies concerning the importance of various demographic factors for changes in aptitude and achievement, both within a cohort of students tested at different ages and between different cohorts tested at the same age-levels.

Within this project six follow-up studies have started.

The first data collection was made in 1961 among all pupils in Sweden born on the 5th, 15th and 25th of any month in 1948. This information for about one tenth of the age cohort was supplemented by data each year as long as the individuals were attending an educational institution. The sample includes a total of some 12 000 individuals, about ninety percent of whom were in the sixth grade within the compulsory school system on the first occasion when data were collected.

In 1966 a new follow-up study was started among all pupils born on the 5th, 15th and 25th of any month in 1953. This sample includes about 10 500 individuals.

The third data collection started in 1980, but now a two-step sampling design was used. First, a stratified sample of 29 out of all 277 local communities in Sweden was drawn. From these communities a systematic sample of classes from grade six was selected. The investigation group consists of the approx. 9 000 students belonging to these classes. This is close to ten per cent of the age-group in the compulsory school. Most of them were born in 1967.

The fourth, the fifth and the sixth samples were drawn in the same way as the third one, but with one exception - the first data were collected as early as in grade 3. Most of the individuals includes in these three samples are borne in 1972, 1977 and 1982 respectively.

Two main categories of data are collected:

Administrative data from school offices and *Project data* obtained directly from the students. Among the latter are results on identical tests - a verbal, a spatial and a reasoning one - given in grade six.

In figure 1 we summarise the design of the project. More information may be found in Härnqvist et al. (1991).

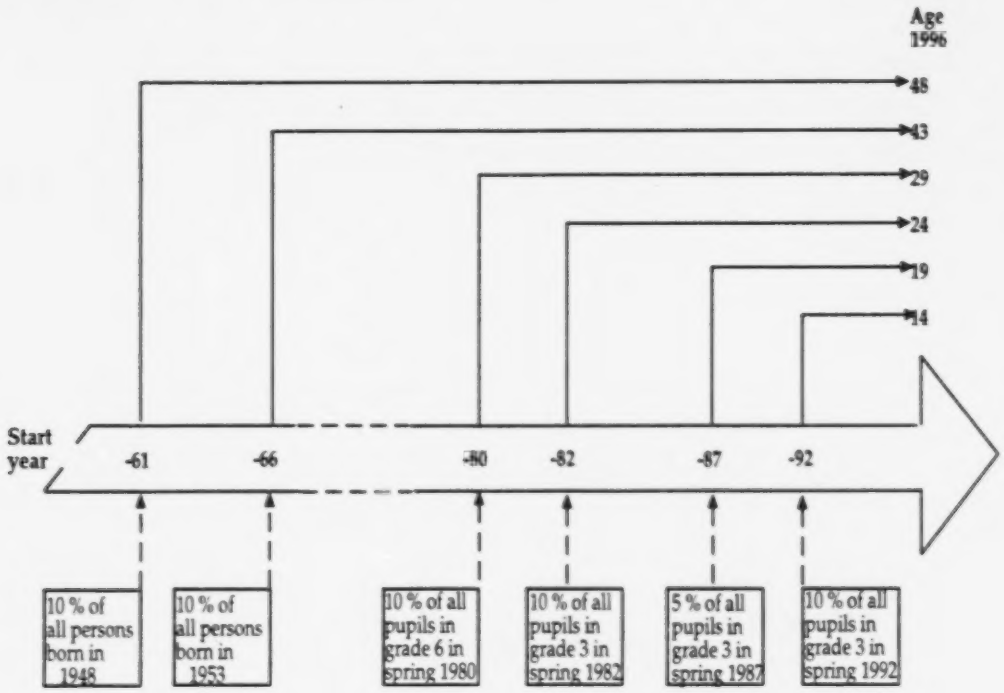

Figure 1. The cohorts included in the Evaluation Through Follow-up project.

Comparisons

The results will be presented in two steps. In step 1 comparisons are made between thirteen-year-old students tested in 1961, 1966 and 1980 and in step 2 comparisons are made between students tested in 1980, 1985, 1990 and 1995. Variations in both composition of the investigation groups and in classification of social background have led to this presentation.

Due to the different techniques of sampling, it was necessary to restrict the first comparison to *students of normal age for grade six*, i.e. 13 years old. This restriction reduces the size of the samples, more among students tested in 1961 and 1966 than in 1980, as the "normal-age" pupils is less than 90 per cent in the first two groups but more than 95 per cent in the third

When comparing students tested in 1980, 1985, 1990 and 1995, it did not seem necessary to restrict the study to students who were in grades appropriate to their age level, because approximately 97 per cent in each

sample met this requirement. Therefore, in step 2 *all students* in grade six are included.

As mentioned previously, the classifications of social background differ too. In step 1, we used a dichotomous scale based on fathers' occupation only. More precisely, we made a rather rough division into two groups called "working-class" and "middle-class" (pupils whose fathers were labourers, and pupils whose fathers were professionals, civil servants, etc.).

In step 2, on the other hand, we used a division into seven groups based on both parents' occupations (see Table IV). This classification was not possible to apply in the first study, as the great majority of the mothers in the two oldest samples were not employed in jobs outside the home.

The sample sizes in the two steps are shown in Table I and II. These tables also report the sample reductions.

"Drop-out group 1" consists of pupils without scores on intelligence tests. In most cases, absence from school on the days of testing accounts for these drop-outs. There is no reason to suspect that these pupils differ in any significant way from the pupils included in the investigations.

"Drop-out group 2", consisting of pupils for whom there is no information about their social background, is found in the first study only. The reason is that in the second study those pupils are included in the analyses - they are placed in a category called group 0.

Table I. Students included in the investigation and drop-outs in the first step.

	Students tested in					
	1961		1966		1980	
	N	%	N	%	N	%
Students included in the investigation	9089	85	8506	88	6515	75
Drop-outs 1 Intelligence data not available	1043	10	873	9	942	11
Drop-outs 2 Background data not available	548	5	282	3	1250	14
Total number of "normal-age" students in the sample	10680	100	9661	100	8702	100

Table II. Students included in the investigation and drop-outs in the second step.

	Students tested in							
	1980		1985		1990		1995	
	N	%	N	%	N	%	N	%
Students included in the investigation	8066	89	7938	85	3984	90	7533	86
Drop-outs 1 Intelligence data not available	1038	11	1429	15	433	10	1272	14
Total number of students in the sample	9104	100	9367	100	4417	100	8805	100

"Drop-out group 2" is rather small among students tested in 1961 and 1966 but larger among pupils tested in 1980 (Table I). This increment is a consequence of the classification of social background being based on fathers' occupation only and that the number of youngsters not living with their biological father had increased a lot from the 60's to the beginning of the 80's.

Table III and IV show the distributions of pupils according to social background. In both steps a clear trend is found so far as the working class categories are gradually decreasing. This is in consequence of structural changes of the Swedish labour market.

Table III. Social background of the students in the first step.

Group	Fathers' occupation	1961	1966	1980
I	Professional, civil servant etc.	50	50	51
II	Labourer	50	50	49
Total		100	100	100

Table IV. Social background of the students in the second step.

Group	Fathers'/mothers' occupation	1980	1985	1990	1995
1	Academic professions	12	16	17	20
2	Civil servants in higher positions	20	22	23	22
3	Civil servants in lower positions	12	12	12	11
4	Farmers, shop-owners etc.	13	9	9	7
5	Skilled workers	19	17	16	15
6	Unskilled workers	18	17	15	15
0	No information	6	8	8	11
Total		100	100	100	100

Variables

The students in the six cohorts were tested with identical intelligence tests. The tests used in the investigations represent the verbal, spatial, and reasoning factors according to a Thurstonian classification of abilities. The tests are called Opposites, Metal Folding and Number Series.

Opposites: To find the opposite of a given word among four alternatives. 40 items, 10 minutes.

Metal Folding: To find the three-dimensional object among four alternatives that can be made from a flat piece of metal with bending lines marked on the drawing. 40 items, 15 minutes.

Number Series: To complete a number series, of which six numbers are given, with two more numbers. 40 items, 18 minutes.

Some statistical data about the tests are reported in Table V. These data are based on the sample tested in 1961.

Table V. Means, standard deviations and reliabilities of the three tests.

Test	Number of items	Means	SD	r_{tt}
Opposites	40	22.88	6.56	.87
Metal folding	40	21.41	7.05	.88
Number series	40	19.41	7.62	.92

As is shown, all the means are fairly close to the midpoint of the possible score range. The standard deviations are of equal size and the reliability coefficients are approximately 0.90, calculated according to the Kuder-Richardson formula 20.

As shown by Gustafsson (1988) the three tests used in this investigation are all loaded on a general factor. This means that the changes over time which we are going to present are caused by changes in the general intellectual ability as well as those in the specific factors measured by the separate tests.

RESULTS

Changes between 1961 and 1980

Means and standard deviations of the scores for boys and girls from different socioeconomic groups tested in 1961, 1966 and 1980, respectively, are found in Appendix (Table AI). As mentioned before, these data are valid for students of normal age for grade six, i.e. students of the age of 13.

In order to give a comprehensive picture of the trends in Table AI a three-way analysis of variance is performed. The results of this analysis is shown in Table VI. In this table only the p-values are given and the significance level is 0.01.

Table VI. Three-way analysis of test means in relation to cohort (Coh), gender (Sex) and socioeconomic group (Soc), p values.

Effects	Verbal	Spatial	Reasoning
Coh	.00	.00	.00
Soc	.00	.00	.00
Sex	.00	.00	.00
Sex x Coh	.00	.01	.01
Soc x Coh	.16	.38	.35
Sex x Soc	.19	.46	.78
Sex x Soc x Coh	.52	.33	.27

As shown by Table VI all the main effects are significant and the same is true for the interactions between gender and cohort. Thus, we can conclude that there are significant differences on all the tests between cohorts, between boys and girls and between students with different socioeconomic backgrounds. Furthermore, the significant interactions between gender and cohort imply that the gender differences in test scores have changed from 1961 to 1980.

The significant main effects for cohorts indicate that the mean score on each test has changed from 1961 to 1980, and this change is shown in figure 2.

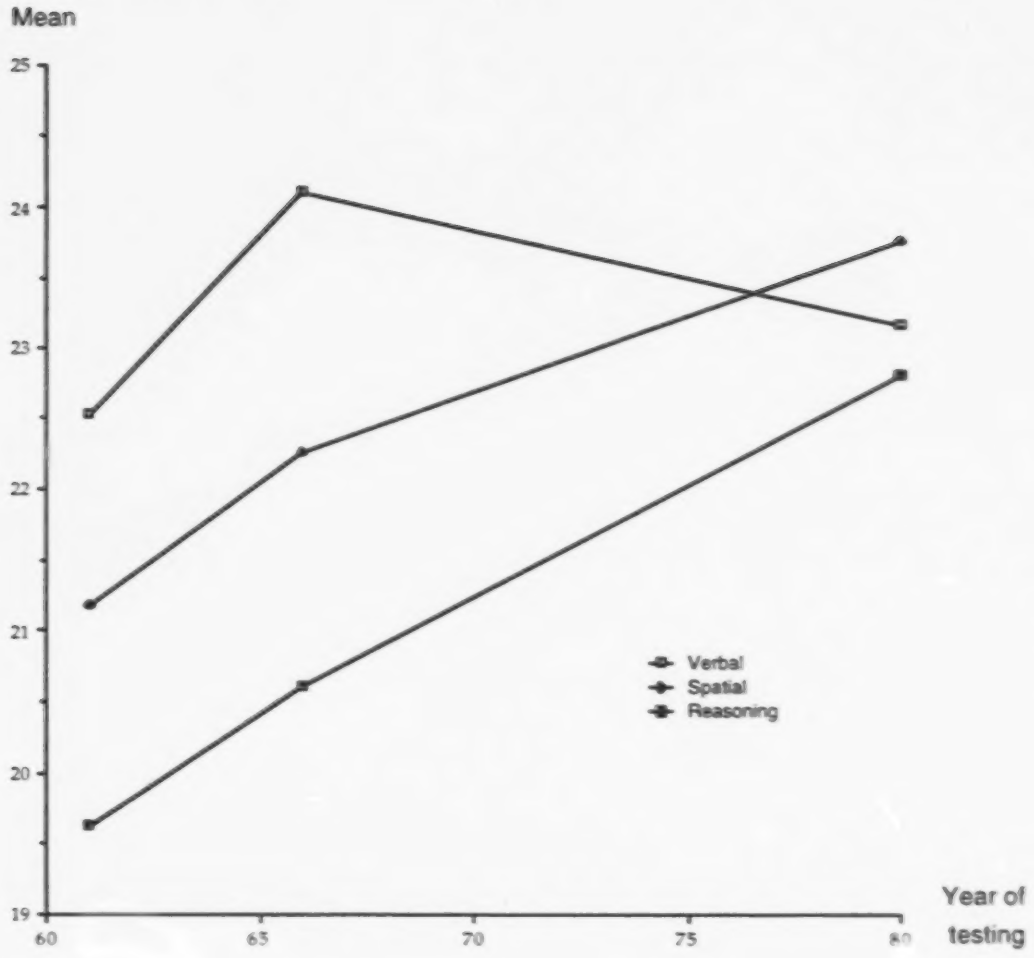

Figure 2. The development of the test mean scores from 1961 to 1980.

All changes are statistically significant which means that there are significant increases in the means on the spatial and reasoning tests between the successive cohorts. The verbal test, on the other hand, shows a different pattern of development. The mean increases between 1961 and 1966, but from 1966 to 1980 there is a significant decline. Still, however, the youngest cohort has performed a higher mean than the oldest one, which means that the overall trend is positive also on the verbal test.

As shown by Table VI there is no significant interaction between cohort and socioeconomic background which means that the development over time is similar for the two socioeconomic groups. Therefore, the actual social differences can be summarized over the cohorts and these differences are shown in Table VII.

Table VII. Socioeconomic differences in test scores. Means.

Group	Verbal	Spatial	Reasoning
High	24.48	23.15	22.18
Low	21.94	21.30	19.37
Difference	2.54*	1.85*	2.81*

* significant at the 1% level

In the traditional way, students from higher socioeconomic groups surpass the others. In this case the differences amount to 2 - 3 points. These differences are greatest on the verbal and spatial tests where they amount to more than one third of a standard deviation. The difference on the spatial test is somewhat less than two raw score points which corresponds to a quarter of a standard deviation. Thus, the results indicate that socioeconomic background has a greater influence on verbal and reasoning abilities than on spatial ability.

The gender differences, on the other hand, have changed over time as was shown by the significant interaction between gender and cohort. Therefore, they have to be presented separately for each cohort.

Table VIII. Gender differences on test scores by cohort. Means.

Test	Year of testing	Girls	Boys	Difference Girls - Boys
Verbal	1961	22.53	22.52	+ 0.01
	1966	24.15	23.96	+ 0.19
	1980	23.40	22.69	+ 0.71*
Spatial	1961	20.34	22.02	- 1.68*
	1966	21.65	22.80	- 1.15*
	1980	23.54	23.69	- 0.15
Reasoning	1961	19.18	19.98	- 0.80*
	1966	20.33	20.76	- 0.43
	1980	22.50	22.51	- 0.01

* significant at the 1% level

As shown by Table VIII the gender differences vary with the test under consideration. On one hand we have the verbal test for which there are no significant differences within the two oldest cohorts, but a significant one in favour of the girls within the youngest group. In no case boys obtain a higher mean than girls on this test. On the other hand we have the spatial and the reasoning tests where the girls in no case surpass the boys. Instead, both these tests show significant differences in favour of boys within the oldest cohort. Also among those born five years later there is a difference in favour of boys, even if it is statistically significant only on the spatial test. The difference of 0.43 in favour of boys on the reasoning test, however, is very close to significance with a p-value of 0.012. On the contrary, the differences found within the youngest group are neglectible on both these tests.

Behind the differing gender differences is a very regular trend of a more favourable development for girls than for boys. On the verbal test, where both genders start on the same level, the girls successively gain an advantage, but on the other two tests, where the boys start ahead, their advantage disappears over the period investigated.

The results presented so far can be summarized:

- all the three abilities studied have shown a positive over all development from 1961 to 1980 even if verbal ability seems to have declined somewhat between 1966 and 1980

- this overall trend is valid irrespective of the students' socioeconomic background

- girls show a more positive trend than boys in all the abilities studied

Changes between 1980 and 1995

As mentioned before, the analyses to be performed in this section comprise all students in grade 6 irrespective of age. This means that some students were either older or younger than 13 years in that grade. However, the great majority - 97 per cent - were normal aged. Furthermore, the classification according to socioeconomic background is more elaborate with seven categories, and now the classification is based on both the parent's occupation.

Means and standard deviations of the scores for boys and girls from different socioeconomic groups tested in 1980, 1985, 1990 and 1995, respectively, are found in Appendix (Table AII - AV).

Table IX. Three-way analysis of variance of the test means in relation to cohort (Coh), gender (Sex) and socioeconomic group (Soc). Changes from 1980 to 1995. (p values).

Effects	Verbal	Spatial	Reasoning
Coh	.00	.00	.04
Soc	.00	.00	.00
Sex	.00	.03	.00
Sex x Coh	.11	.10	.07
Soc x Coh	.00	.14	.62
Sex x Soc	.16	.68	.01
Sex x Soc x Coh	.83	.82	.60

All main effects except that for cohort on the reasoning test and that for sex on the spatial test are statistically significant. This means that the verbal and the spatial abilities have continued to change during the period from 1980 to 1995, while there are only minor changes in the reasoning ability. In the same way as for the period from 1961 to 1980 there are significant socioeconomic differences and there are also gender differences in the verbal and reasoning abilities. On the contrary, no significant difference remains between boys and girls in spatial ability.

In the preceding section we found that the gender differences had changed from 1961 to 1980. After 1980 they seem to remain stable, as shown by the non-significant interaction effects between sex and cohort. In stead, there is a significant interaction between socioeconomic background and cohort in verbal ability indicating that the social differences in this ability have changed over time. Furthermore, there is a significant interaction between gender and socioeconomic background in reasoning ability and this implies that the gender differences in this variable vary with socioeconomic background - an interaction which is valid for all cohorts.

The development of the test mean scores from 1980 to 1995 is shown in Figure 3.

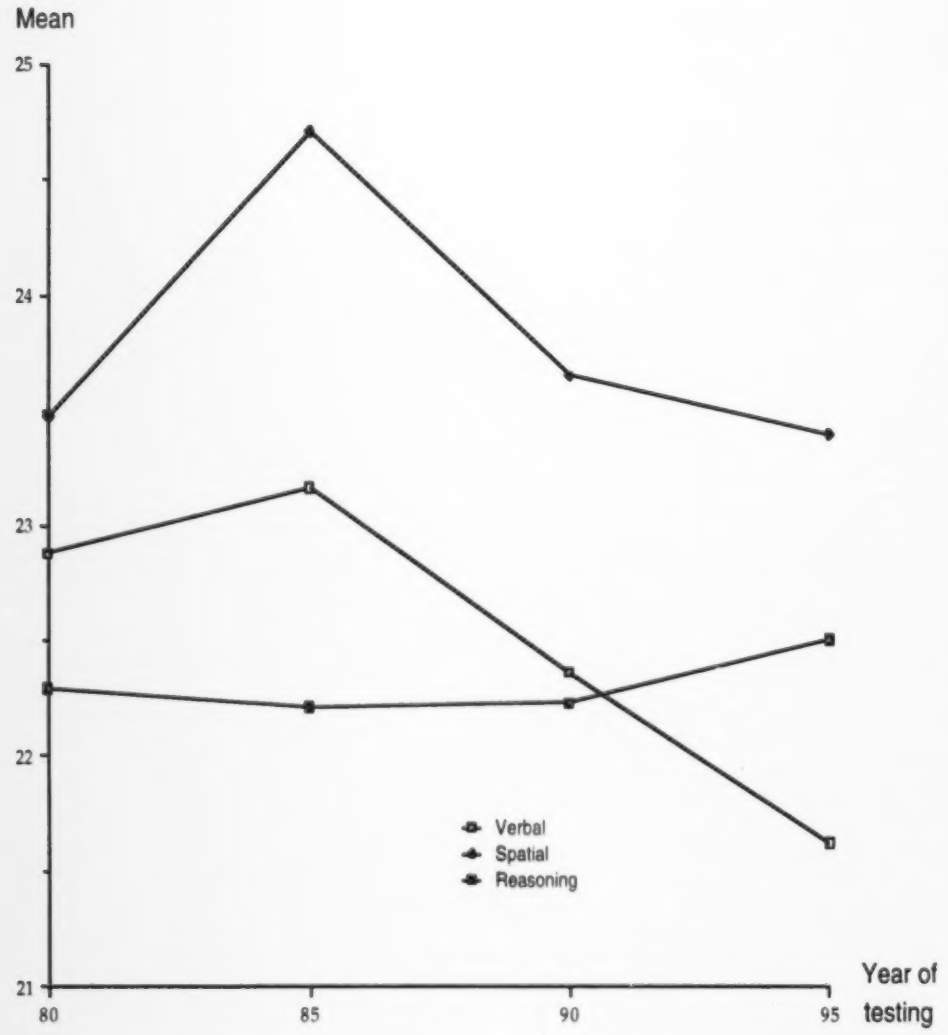

Figure 3. The development of the test mean scores from 1980 to 1995.

From figure 3 we can see that verbal and spatial abilities have changed in a similar but not identical way. They have both increased significantly from 1980 to 1985 but after that they have successively declined. The declination, however, is strongest for the verbal ability so that the mean for the youngest cohort is below that for the oldest one. The weaker declination of the spatial ability implies that the mean for the youngest cohort is about the same as

that for the oldest one (students tested in 1980). Reasoning ability, finally, shows a much less dramatic development and remain on about the same level for all the four cohorts.

Thus, we can conclude that verbal, spatial and reasoning abilities have developed quite differently from 1980 to 1995 as compared to 1961 to 1980. During the last mentioned period they all developed positively and this trend has been stable for verbal and spatial abilities also up to 1985. After that they have both declined. The reasoning ability ceased its positive development already after 1980 and since then it has remained on the same level.

Besides the cohort differences there are also socioeconomic differences in all the three abilities. Those found in verbal ability, furthermore, have changed over time, while the differences in spatial and reasoning abilities have been fairly constant. We start by inspecting the stable socioeconomic differences in spatial and reasoning abilities (Table X).

Table X. Socieconomic group means in spatial and reasoning during the period 1980 to 1995.

	Socioeconomic group						
	1	2	3	4	5	6	0
Spatial	25.99	24.95	24.05	23.61	22.87	22.07	21.95
Reasoning	25.63	23.74	22.22	22.19	20.72	19.98	19.89

The rank order between the socioeconomic groups are exactly the same according to both abilities and it goes from a highest mean for group 1 to a lowest for group 0. All differences are statistically significant except those between groups 3 and 4 and those between groups 6 and 0, respectively. The table also shows that the differences between the highest and the lowest group are substantial. On the spatial test the mean range amounts to 4 raw score points, which corresponds to more than half a standard deviation and on the reasoning test it is even greater - nearly 6 points which is close to one standard deviation unit.

Taking into consideration the results shown in the preceding section, we can conclude that the socioeconomic differences in spatial and reasoning abilities have remained big from the early 60's up to the mid 90's, and there are no signs of a development towards decreasing social differences in these respects, at least not during the latest 15 years.

As the three-way analysis of variance showed, the social differences in verbal ability have not been equally stable from 1980 to 1995. Therefore, we

present the socioeconomic means on this test for each cohort. In order to get a summarizing measure of the total differences we also give the intraclass correlations for each cohort. A high intraclass correlation indicates great socioeconomic differences.

Table XI. The development of socieconomic differences in verbal ability from 1980 to 1995. Means and intraclass correlations (C).

Cohort	Socioeconomic group (Means)							C
	1	2	3	4	5	6	0	
1980	26.24	23.98	22.66	23.18	21.54	21.19	21.78	.070
1985	25.86	24.50	22.51	23.33	21.73	21.18	21.43	.078
1990	24.90	23.38	22.21	22.66	20.82	20.07	20.46	.080
1995	24.40	22.96	21.52	21.45	20.39	19.32	18.54	.111

From the intraclass correlations in the last column of Table XI it is obvious that the social differences in verbal ability has increased over time, and this is primarily due to the big differences found within the youngest cohort. Certainly, a comparison between 1980 and 1990 also shows an increase but it is more moderate.

The three-way analysis of variance showed significant gender differences in verbal and reasoning abilities. Furthermore, there was a significant interaction between socioeconomic background and gender in reasoning ability - an interaction which is similar for all the four cohorts. Therefore, the gender differences and the interaction can be shown for all cohorts combined.

In the preceding section we reported that there were no gender differences at all in verbal ability in 1961. After that, however, the girls had a more positive trend, and they surpassed the boys by about 0.7 raw score points in 1980 - a difference which is statistically significant. Also the younger cohorts show gender differences in the same direction. Taken over all the four younger groups the girls' mean amounts to 22.88 which is to be compared to 22.22 for the boys.

In reasoning ability the previous analyses showed that the boys surpassed the girls within the oldest cohort, but this difference had disappeared up to 1980. As can be seen in Table XII there is again a gender difference in favour of boys within the four youngest cohorts and this difference of 0.3 raw score points is statistically significant.

Furthermore, the three-way analysis of variance showed that the gender differences in this variable varied with socioeconomic background. From Table XII we can see that this is due to significant differences in favour of boys within socioeconomic group 1 and 4, i.e. among those children whose parents belong to the academic professions and to farmers, shop owners etc. On the other hand there is quite a substantial difference - although not statistically significant - in favour of girls within group 3 which contains civil servants in lower positions.

Table XII. Gender differences in reasoning ability by socioeconomic background. 1980 - 1995.

	Socioeconomic group							All
	1	2	3	4	5	6	0	
Boys	26.06	23.85	21.90	22.57	20.93	20.00	20.15	22.49
Girls	25.18	23.62	22.54	21.81	20.50	19.96	19.63	22.14
Diff.	0.88*	0.23	-0.64	0.76*	0.43	0.04	0.52	0.34*

* significant at the 1% level

DISCUSSION

The development from 1960 through 1980
If we compare the interval 1961 - 66 with 1966 - 80, the former seems more favourable because during this period the means of all the tests are rising, and all the increases are rather large in spite of the shorter time interval.

One of the significant causes of the positive change during the first half of the 60's, especially for female students, may be the introduction of a new school organisation. Most of the students tested in 1961 attended elementary school (folkskola) and a smaller part a so-called experimental comprehensive school (enhetsskola). In 1966 the majority of students attended the comprehensive school (grundskola) established according to the new school law from 1962. It is difficult to explain how the differences in organisation and curriculum may have influenced the behaviour of the students, and thereby their reactions to the tests. However, that the school reform has been of importance, is supported by the fact that a statistically significant correlation has been reported earlier between the improvement in girls' spatial ability and the introduction of the new comprehensive school (Härnqvist & Stahle, 1977).

Looking at the period 1966 to 1980, the trends are not as positive, mostly due to decreasing results on the verbal test for all groups. However, this finding probably depends more on some special characteristics of the test than on a decreasing verbal ability among the students. This interpretation is supported by the fact that some items show decreasing proportions of correct answers at the same time as others show increasing proportions. Furthermore, there is a systematic pattern in the changes on item-level. Most of the items with increasing proportions may be classified as foreign words. Those with a decreasing proportions of correct answers are very often elderly words, probably used more in everyday talk and texts in the 1960's than at the beginning of the 80's. This can be said to illustrate one big difficulty in testing changes in verbal ability over longer periods. The difficulty of items in a vocabulary test, will increase or decrease along with changes in word frequencies in everyday language. Therefore, test results from different occasions may not be completely comparable, even if, as in our study, identical tests are used for every test run.

The means of the spatial and reasoning tests also show a continuous increase from 1966 to 1980. However, considering that this period is almost three times as long as the previous one (1961-66), these increments are of relatively moderate sizes. One reason may be that there is no longer a "push off" from what is learned in the comprehensive school - because most of the students tested in 1966 already attended this kind of school.

To explain the increments of spatial and reasoning ability between 1966 and 1980 other explanations have to be sought. One may be pre-school activities.

The students tested in 1980 had taken part in such activities to a much greater extent than those tested in 1966. Experiences of pre-school activities where problem solving and experimenting are stressed, are likely to have contributed to an increase in ability to solve test items in the spatial and reasoning tests (cf. Andersson, 1992).

The development from 1980 through 1995

In the verbal test results there is a weak rising trend between 1980 and 85, followed by a somewhat stronger declining trend between 1985 and 95. However, as mentioned previously, it is difficult to interpret changes over time in a vocabulary test, due to the fact that the specific items included in the test become more or less frequent in everyday language.

In the spatial test, there is a substantial rise of the means during the first half of the 80´s, followed by a decline of almost the same size. What explanations are to be found? Let us look first at the large rise between 1980 and 1985. During this five-year period the increases are larger than between 1961 and 1966, when the great reform of the compulsory school contributed to the improved results in the spatial test (Härnqvist & Stahle, 1977). There is no self-evident explanation to be found for the change in results between 1980 and 1985. However, what has to be taken in account, is the fact that the students tested in grade 6 in 1985, were tested also in grade 3 with a "metal folding test". This has not been the case for any other cohort. For that reason, we assume that the extremely good results in the spatial test in 1985, to some extent at least, is a consequence of test-training. Both the unique character of the test and the fact that the increments have disappeared in 1990, speaks in favour of such an interpretation

The differences in the reasoning test results are very moderate from 1980 to 1995 and the magnitude of almost all the changes are of such a size that they may be explained by sampling errors.

Changes in gender differences

Although the trends are similar in character for male and female students over the whole 35-year period, we have found some differences in the amounts of change. This is very clearly seen in Table XIII.

In the verbal test, gender differences were very small and insignificant in 1961 but in 1980 the girls show a significantly higher mean. In spatial and reasoning ability, the boys were superior at the beginning of the sixties - 20 years later this superiority had disappeared completely. Thus, there has been a more positive trend for girls than for boys during the 60's and the 70's. From the data already given, we can conclude that the changes since 1980 have not favoured girls in the same manner as during the two previous decades. This is still more obvious if we look at table XIII. In verbal and spatial ability they keep their gains from 1960 to 1980 but they do not extend them. In reasoning ability on the other hand they tend to loose a great deal of what they have attained during the 60's and 70's.

Table XIII. Gender differences in the three tests. Positive differences indicate a higher mean for female students.

	Test		
	Verbal	Spatial	Reasoning
1961	+0.01	-1.68*	-0.80*
1980	+0.71*	-0.15	-0.01
1995	+0.63*	-0.36	-0.57*

* significant at the 1% level

Nevertheless, if we watch the whole period from 1961 to 1995 it is obvious that female student changes stand out in a somewhat more favourable way than those of the male students. At the beginning of the period the boys were superior in spatial and reasoning ability and there were no gender differences in verbal ability. 35 years later the girls had an advantage in verbal ability and the boys achieved better on the reasoning test only.

However, it must be kept in mind that the gender differences in every cohort and in every test are rather small compared to the differences between students from different social backgrounds - which we are going to discuss in the following section.

Changes in socioeconomic differences

The differences between students from various socioeconomic groups, based on the results from the first step of the study, are found in Table XIV and those from the second step in Table XV. Most of the differences are considerably larger in the latter table. This is a consequence of the transformation from a dichotomous scale to a classification consisting of seven groups.

Table XIV. Socioeconomic differences in the three tests during the period from 1961 through 1980.

	Test		
	Verbal	Spatial	Reasoning
Soc. diff. I-II	2.54*	1.85*	2.81*

* significant at the 1% level

Table XV. Socioeconomic differences in the three tests during the period 1980 through 1995. Differences between group 1 and the other groups.

Test		Socioeconomic group					
		2	3	4	5	6	0
Verbal	1980	2.26*	3.58*	3.06*	4.70*	5.05*	4.46*
	1995	1.44*	2.88*	2.95*	4.01*	5.08*	5.86*
Spatial	1980-95†	1.04*	1.94*	2.38*	3.12*	3.92*	4.04*
Reason.	1980-95†	1.89*	3.41*	3.44*	4.91*	5.65*	5.74*

† Average difference for the period
* significant at the 1% level

The socioeconomic differences show a greater stability than the gender differences. Over the whole period of almost 35 years the socioeconomic differences in spatial and reasoning ability have remained remarkably stable, and the only statistically significant change which has occurred is an increase of the differences in verbal ability. The last mentioned change is most striking between the two youngest cohorts (students tested in 1990 and 1995 respectively) where particularly group 0 seems to fall behind the other groups. This may be an indication of increasing socioeconomic differences at least to some extent may be due to a growing number of immigrant youngsters, as children from immigrant homes are strongly over-represented in group 0.

The increasing differences in verbal ability between students from different socioeconomic groups must be judged as something very serious. Bearing in mind that the test scores not only tell us about the development of verbal ability up to the age of thirteen, but also have significant implications for future success in school and higher education (Svensson, 1971; Reuterberg et al., 1993; Härnqvist, 1994), this observed trend has to be seen as important. It may be an indication of still growing gaps in the future between children with different economic and cultural backgrounds. If so, it is an educational challenge, since such a trend threatens the possibilities of reaching the equality goals expressed in the school curricula.

SUMMARY

In Table XVI, a review of the changes in test scores from 1961 to 1995, separated into the periods before and after 1980, is given.

The verbal test scores tend to rise up to 1980, whereupon there is a stronger tendency in the opposite direction. The consequence of these contradictory trends is that the pupils' verbal ability seems to be lower in 1995 than it was thirty-five years earlier. However, behind this decrease, big differences on item-level are hidden. The proportion of correct answers among older (archaic) words has diminished and the proportion among foreign words has grown, which is rather obvious considering the continuous development of languages. These conditions make it difficult to talk about changes in verbal ability over longer periods, especially if this ability is measured by a vocabulary test. The only conclusion that may be drawn is that the vocabulary of the students is partly different nowadays than it was thirtyfive years earlier.

The results in the spatial test have risen considerably up to 1995. For the girls there is an increase of almost half a standard deviation unit, and for the boys about a third of a unit since 1961. However, the whole of the girls' increment, and the largest part of that for the boys, has taken place during the 60's and the 70's. Among the factors which may have caused this development is the introduction of the comprehensive school, and the growth of the pre-school - two circumstances which have probably contributed to both a rising spatial ability and an equalization of the sex differences in this ability.

In the reasoning test, the increments for both sexes are of the same magnitude as in the spatial test for girls, and as in that test, almost all of the increments are assigned to the first two decades of the period. Even in this case school reforms are likely to have been of importance.

Table XVI. Changes in test results between 1961 and 1995

Test	Difference	Boys	Girls
Verbal	1980 - 61	+ 0.17	+ 0.87
	1995 - 80	- 1.16	- 1.36
	1995 - 61	- 0.99	- 0.49
Spatial	1980 - 61	+ 1.67	+3.20
	1995 - 80	+ 0 04	- 0.21
	1995 - 61	+ 1.71	+ 2.99
Reasoning	1980 - 61	+ 2.53	+ 3.32
	1995 - 80	+ 0.44	- 0.12
	1995 - 61	+ 2.97	+ 3.20

REFERENCES

Andersson, B-E (1992). Effects of day-care on cognitive and socioemotional competence of thirteen-year-old Swedish schoolchildren. *Child Development, 63,* 20-36.

Brody, N. (1992). *Intelligence.* 2nd ed. San Diego: Academic Press.

Emanuelsson, I., Reuterberg, S-E, & Svensson, A. (1993). Changing differences in Intelligence? *Scandinavian Journal of Educational Research, 37,* 259- 277.

Flanagan, J. C. (1976). Changes in school levels of achievement. *Educational Researcher, 5,* 9-12.

Flynn, J. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin, 101,* 171-191.

Gustafsson, J-E (1988). Hierarchial models of individual differences in cognitive abilities. In R. J. Sternberg, (Ed.) *Advances in the Psychology of Human Intelligence, 4,* 35-71.

Halpern, D. F. (1992). *Sex Differences in Cognitive Abilities.* 2nd ed. New Jersey: Lawerence Erlbaum Associates.

Harnischfeger, A., & Wiley, D. E. (1976). Achievement test scores drop. So What? *Educational Researcher, 5,* 5-12

Husén, T., & Tuijnman, A. (1991). The cotribution of formal schooling to the increase in intellectual capital. *Educational Researcher, 20,* 17-25.

Härnqvist, K. (1968). Relative changes in intelligence from 13 to 18. *Scandinavian Journal of Psychology, 9,* 50-82.

Härnqvist, K. (1994). Social selektion till gymnasieskola och högskola. (Social selection to upper secondary school and higher education). I R. Erikson, & J. Jonsson, *Sorteringen i skolan.* Stockholm: Carlssons bokförlag.

Härnqvist, K., Emanuelsson, I., Reuterberg, S-E, Rosén, M., & Svensson, A. (1991). Evaluation through follow-up. A longitudinal Swedish project. Department of Education. University of Gothenburg.

Härnqvist, K. & Stahle, G. (1977). An ecological analysis of test score changes over time. *Reports from the Institute of Education. University of Gothenburg.* Nr 64.

Lynn, R. (1990). The role of nutrition in secular increases in intelligence. *Personality and Individual Differences, 11,* 273-285.

Reuterberg, S-E, Emanuelsson, I., & Svensson, A. (1993). Changes in verbal and mathematical achievement. Paper presented at the 5th EARLI conference, Aix-en Provence, August 31 - September 5, 1993.

Svensson, A. (1971). *Relative achievement: School performance in relation to intelligence, sex and home environment.* Stockholm: Almqvist & Wiksell.

APPENDIX

Table A I. Means and standard deviations of the tests. Students tested in 1961, 1966, and 1980, respectively, categorized according to gender and socioeconomic group.

			Group I		Group II	
Test	Gender	Cohort	M	SD	M	SD
Verbal	Male	1961	23.72	6.77	21.27	6.42
		1966	25.25	6.31	22.74	6.20
		1980	23.84	5.63	21.45	5.61
	Female	1961	24.06	7.00	21.17	6.54
		1966	25.35	6.53	22.98	6.30
		1980	24.61	5.79	22.14	5.86
Spatial	Male	1961	22.93	7.22	21.07	7.51
		1966	23.69	7.40	21.91	7.39
		1980	24.69	7.23	22.70	7.61
	Female	1961	21.27	6.76	19.53	6.77
		1966	22.30	6.83	21.06	7.04
		1980	24.49	6.69	22.60	7.07
Reasoning	Male	1961	21.23	7.70	18.81	7.87
		1966	22.13	7.84	19.42	7.89
		1980	24.11	8.08	21.08	8.16
	Female	1961	20.68	7.52	17.88	7.61
		1966	21.56	7.67	19.17	7.52
		1980	24.01	7.65	21.23	7.74

Table A II. Means and standard deviations of the tests. Students tested in 1980 categorized according to gender and socioeconomic group.

		Boys		Girls	
Test	Soc	M	SD	M	SD
Verbal	1	25.94	5.73	26.53	5.69
	2	23.60	5.42	24.37	5.78
	3	22.01	5.72	23.32	5.75
	4	22.69	5.72	23.67	5.96
	5	21.18	5.78	21.94	5.80
	6	20.90	5.49	21.49	6.06
	0	21.45	5.62	22.07	6.72
Spatial	1	25.74	7.37	25.94	6.17
	2	24.70	7.15	24.56	6.98
	3	23.87	7.41	23.74	6.63
	4	23.19	7.44	23.09	6.88
	5	22.68	7.75	22.81	7.07
	6	22.08	7.60	21.65	7.28
	0	22.26	7.86	21.85	7.43
Reasoning	1	26.35	7.97	25.54	7.33
	2	23.76	8.01	23.73	7.76
	3	22.23	8.44	22.99	7.70
	4	22.23	8.42	22.18	7.98
	5	20.67	8.18	20.98	7.80
	6	20.08	8.19	20.49	7.82
	0	20.45	8.74	20.08	7.78

Table A III. Means and standard deviations of the tests. Students tested in 1985 categorized according to gender and socioeconomic group.

		Boys		Girls	
Test	Soc	M	SD	M	SD
Verbal	1	25.48	5.42	26.26	5.84
	2	24.20	5.34	24.84	5.92
	3	21.82	5.78	23.21	5.48
	4	22.94	5.43	23.75	6.13
	5	21.51	5.75	21.94	5.89
	6	20.74	6.07	21.65	6.14
	0	21.13	6.48	21.72	6.28
Spatial	1	26.68	7.08	26.61	6.08
	2	26.03	7.57	25.79	6.67
	3	24.66	7.64	25.31	6.75
	4	24.44	7.87	24.67	6.70
	5	23.62	7.94	23.52	7.34
	6	22.98	7.95	22.86	7.62
	0	23.00	8.32	23.67	7.17
Reasoning	1	25.95	7.90	24.88	7.65
	2	23.91	8.35	23.37	7.59
	3	21.38	8.64	22.27	7.48
	4	21.91	8.16	21.84	7.66
	5	20.72	8.49	20.49	7.98
	6	20.13	8.47	19.89	7.91
	0	20.58	9.13	19.94	8.26

Table A IV. Means and standard deviations of the tests. Students tested in 1990 categorized according to gender and socioeconomic group.

Test	Soc	Boys M	SD	Girls M	SD
Verbal	1	24.91	5.21	24.89	5.83
	2	23.12	5.45	23.68	5.50
	3	22.05	5.81	22.37	5.49
	4	22.47	5.60	22.86	5.69
	5	20.94	5.22	20.69	5.78
	6	19.84	5.74	20.31	6.16
	0	20.26	6.58	20.67	6.38
Spatial	1	25.62	7.31	25.49	6.27
	2	24.83	7.18	24.18	6.94
	3	24.23	7.68	23.66	7.26
	4	23.97	7.44	22.78	6.76
	5	22.65	7.83	21.71	7.09
	6	22.43	7.92	22.16	7.55
	0	21.87	8.11	22.18	8.38
Reasoning	1	25.48	8.03	24.97	7.65
	2	23.72	8.53	23.25	7.94
	3	22.54	8.64	22.11	8.40
	4	23.38	8.31	21.18	7.74
	5	21.00	8.53	19.74	8.11
	6	19.59	8.50	19.92	8.48
	0	20.03	9.55	18.88	9.67

Table A V. Means and standard deviations of the tests. Students tested in 1980 categorized according to gender and socioeconomic group.

Test	Soc	Boys M	SD	Girls M	SD
Verbal	1	23.98	5.21	24.82	5.42
	2	22.50	5.25	23.48	5.58
	3	20.78	5.66	22.23	5.72
	4	21.08	5.16	21.84	5.25
	5	20.14	5.37	20.62	5.60
	6	19.46	5.88	19.17	5.66
	0	18.49	6.18	18.60	6.29
Spatial	1	25.99	7.35	25.46	6.67
	2	24.53	7.63	24.49	6.67
	3	23.34	7.40	23.42	6.81
	4	23.97	7.70	22.50	7.13
	5	22.78	7.62	22.39	7.12
	6	21.14	7.89	21.18	7.38
	0	21.04	8.15	20.72	7.89
Reasoning	1	26.27	7.95	25.32	7.53
	2	23.94	8.85	23.97	7.54
	3	21.49	8.84	22.31	8.04
	4	23.28	8.69	21.69	8.18
	5	21.51	8.69	20.40	8.17
	6	19.96	9.26	19.36	7.76
	0	19.76	9.44	19.41	8.63